AMENDMENTS TO THE SPECIFICATION

The following paragraph is being added beginning on page 6, after line 21:

Fig. 2 is a flowchart of a text generation method in accordance with the present invention.

The paragraph beginning on page 6, line 22, is being amended as follows:

~~Fig. 2~~ Fig. 3 is a subgraph illustrating a dependency structure analyzed by a text generation unit.

The paragraph beginning on page 6, line 24, is being amended as follows:

~~Fig. 3~~ Fig. 4 is a dependency tree generated by the text generation unit.

The paragraph beginning on page 7, line 1, is being amended as follows:

~~Fig. 4~~ Fig. 5 is a dependency tree in another sample sentence.

The paragraph beginning on page 7, line 2, is being amended as follows:

~~Fig. 5~~ Fig. 6 illustrates an example of calculation of a probability that an order of word dependency is appropriate.

The following paragraph is being added beginning on page 7, after line 3:

Fig. 7 is a flowchart of a word insertion process in accordance with the present invention.

The paragraph beginning on page 7, line 24, is being amended as follows:

If the keyword input unit (10) inputs three keywords (2) of "kanojo", "kouen", and "itta", the text and phrase searching and extracting unit (11) searches and extracts a text or a phrase, each containing at least one of the keywords captured from the database (13).

The paragraph beginning on page 8, line 7, is being amended as follows:

This process will be discussed in more detail. ~~In response to the keyword input by the keyword input unit (10), the text and phrase searching and extracting unit (11) extracts a sentence having n keywords from the database (13). It is perfectly acceptable if one keyword is contained in the sentence. The extracted sentence is then sent to the text generation unit (12).~~ Fig. 2 illustrates a flowchart of a text generation method. The keyword input unit (10) performs a keyword capturing process (20) for capturing into the text generation apparatus (1) a keyword that is input by any input means such as an unshown keyboard, an unshown touch panel input device, or an input device of another computer connected to a network.

The following paragraphs are being added beginning on page 8, after line 13:

The text and phrase searching and extracting unit (11) searches the database (13) for a sentence containing n keywords (21) in response to a keyword (2) captured by the keyword input unit (10), and extracts a found sentence (22). Searching techniques for searching the database (13) for a sentence matching a keyword are known as a character string extraction method, and a technique permitting high-speed extraction is preferably used.

A vast amount of sentences is available in the internet, and these contents can be used as the database (13). Searching for contents may be limited to particular sites. For example, when a text relating to current events is generated, the use of sites of newspaper publishers may contribute to the generation of a precise text.

It is sufficient if one keyword is contained at any rate. The extracted sentence is sent to the text generation unit (12). If no sentence is found in the text and phrase searching process (21), the text generation apparatus (1) requests a user to add another keyword (23), and an additional keyword is input using the keyword input unit (10).

The paragraph beginning on page 8, line 14, is being amended as follows:

The text generation unit (12) includes the parser (12a), the constructor (12b), and the evaluator (12c). The parser (12a)

performs a morphologically ~~analyzes and parses~~ analysis process (24) and a parsing process (25) on the extracted sentence.

The paragraph beginning on page 8, line 18, is being amended as follows:

Available as a morphological analyzing method (24) is a method of analyzing a morpheme based on an ME model, as disclosed in Japanese Patent Application No. 2001-139563 applied by the applicant of this application.

The paragraph beginning on page 10, line 7, is being amended as follows:

A parsing method (25) ~~using an ME model may be used as a parsing method~~ of the parser (12a) may also include a parsing technique based on an ME model. Any other parsing method may be used. The following method is used in one embodiment. The text generation unit (12) may references the database (13), and learns a plurality of texts contained the database (13) in the ME model.

The paragraph beginning on page 12, line 24, is being amended as follows:

Subgraphs are searched (21) and extracted (22) from the database (13) according to the above keywords (2) and are analyzed. The subgraphs having high frequencies of occurrence are ones illustrated as a and b in Fig. 3a.

Fig. 3a shows a sentence having a dependency structure of "<noun> wa <noun> e <verb>". Fig. 3b shows a sentence having a

dependency structure of "<noun> no <noun> e <verb>". These sentences have respective keywords.

~~Figs. 2a and 2b illustrate the subgraphs having high frequencies of occurrence.~~ Referring to Fig. ~~2a~~ 3a, the keyword (kanojo wa) is a node (parent node 1) ~~(20)~~ (30), and "<noun>+e" is a node (parent node 2) ~~(21)~~ (31), and "<verb>." is a node (child node) ~~(22)~~ (32), and a dependency ~~relation (23)~~ structure (33) results.

The paragraph beginning on page 13, line 11, is being amended as follows:

It is assumed that n input keywords are in a dependency relation, and a dependency structure tree containing the n input keywords is generated (26). To generate the tree, the subgraphs are combined.

The paragraph beginning on page 13, line 15, is being amended as follows:

For example, the three keywords are input, and it is assumed that the three keywords are in a dependency relation, and the subgraphs are combined (in this case, aligned). Trees shown in Figs. ~~3a and 3b~~ 4a and 4b thus result.

The paragraph beginning on page 13, line 19, is being amended as follows:

The above-referenced parsing method (dependency model) is again used to select which of the two generated trees (Figs. ~~3a~~

~~and 3b~~ Figs. 4a and 4b) as appropriate in an ordering process (26a).

The paragraph beginning on page 13, line 22, is being amended as follows:

For the ordering of the dependency tree (26a), the ratio of agreement between a combination of subgraphs, the frequency of occurrence, and the dependency relation are taken into consideration. If n is three or more, an ambiguity is present in the dependency relation between the n words. To solve the ambiguity, a dependency model is used. A word having a larger probability determined from the dependency model is ordered with higher priority.

The paragraph beginning on page 14, line 5, is being amended as follows:

As a result, the probability of the tree of Fig. 3a is higher, and the tree of Fig. 3a is selected as the optimum dependency ~~relation~~ structure (27).

The paragraph beginning on page 14, line 9, is being amended as follows:

To contribute to the output of a more natural text in Japanese language, the most natural word order is preferably selected. In accordance with the present invention, the following re-arrangement of word order is possible. (Word order determination process (28))

The paragraph beginning on page 14, line 18, is being amended as follows:

From the tree having the higher priority, a sentence is re-arranged in the natural word order and is output. Used to this end is a word order model based on the ME model that generates a natural order sentence from a dependency structure (28a). The database (13) may be referenced to learn the word order model.

The paragraph beginning on page 15, line 13, is being amended as follows:

The embodiment of the present invention provides a technique to learn a relationship between elements in a sentence and the tendency of word order, namely, a regularity from a predetermined text. In the determination of word order (28), this ~~This~~ technique learns the word order by referring to what element contributes to the determination of word order in what degree but also what combination of the elements results what tendency of the word order. This technique thus deductively learns a text. The degree of contribution of each element is efficiently learned using the ME model. The word order is learned by sampling two phrases at a time regardless of the number of modified phrases.

The paragraph beginning on page 15, line 25, is being amended as follows:

To generate a sentence, the learned model is used. With the phrases in dependency relation received, the order of the dependency phrases are determined. The decision of the word

order (28) is performed as below.

The paragraph beginning on page 16, line 18, is being amended as follows:

For example, an optimum word order is now determined in a sentence "kinou (yesterday)/tenisu wo (tennis)/Taro wa (personal name)/shita (played)." In the same way as already discussed, a dependency tree is produced. A structure tree having the highest probability is obtained as shown in ~~Fig. 4~~ Fig. 5.

The paragraph beginning on page 16, line 24, is being amended as follows:

More specifically, words modifying verb "shita." (43) include three namely, "kinou" (40), "tenisu wo" (41), and "Taro wa" (42). The order of the three words are determined (29).

The paragraph beginning on page 17, line 2, is being amended as follows:

~~Fig. 5~~ Fig. 6 illustrates a calculation example (50) of a probability that the order of the dependency phrases is appropriate.

The paragraph beginning on page 17, line 9, is being amended as follows:

For example, the probability of the word order of "kinou" and "Taro wa" in the chart is "p*(kinou, Taro wa)", and is assumed to be 0.6. Similarly, the word order of "kinou" and

"tenisu wo" is 0.8, and the word order of "Taro wa" and "tenisu wo" is 0.7, and the probability of the word order (51) at a first row in ~~Fig. 5~~ Fig. 6 is determined by multiplying the probabilities, and is thus 0.336.

The paragraph beginning on page 17, line 25, is being amended as follows:

If a generalized node is contained in the word order model (28a), the node is presented as is, and a location where a personal name, a geographic name, or a date is easy to place is known.

The paragraph beginning on page 18, line 4, is being amended as follows:

The dependency structure is received in the word order model (28a) in the above-referenced word order model. In accordance with the embodiment of the present invention, a word order model is used in a building process of the dependency structure.

The paragraph beginning on page 18, line 18, is being amended as follows:

The evaluator (12c) evaluates the text candidates (29) by putting together various information including the order of the input keywords, the frequency of occurrence of the extracted pattern, and a score calculated from the dependency model and the word order model. The evaluator (12c) may reference the database (13).

The paragraph beginning on page 20, line 13, is being amended as follows:

The present invention provides an insertion method that is performed when keywords are not sufficient. Fig. 7 is a flowchart of an insertion process.

The paragraph beginning on page 20, line 15, is being amended as follows:

If n keywords are input (60), inter-word space is filled using the ME model. Two keywords out of n keywords are extracted (61) and input to the model, and the insertion process is performed between the two keywords.

The paragraph beginning on page 20, line 19, is being amended as follows:

A determination is made of whether there is a word to be inserted between any two keywords (62). If the number of words to be inserted is one word, that word is inserted (65). If there ~~are~~ is a plurality of words to be inserted between the two keywords (63), the probability of occurrence of each of the words is determined (64). An insertion operation is performed starting with a word having the highest probability (65). This process is performed for each of any two words (66).

The paragraph beginning on page 21, line 1, is being amended as follows:

The insertion operation is terminated when the probability

of "no insertion" becomes highest between any two keywords (67).